riivo

AI-NATIVE PRODUCT BUILD · CUSTOM ENGINEERING

PROPOSAL FOR CHRISTY KIRKPATRICK · EMPANDA

AI Assessment and Marking Platform

A concept-proving MVP that marks qualitative answers instantly, consistently and defensibly, with an accredited assessor overseeing every release.

01 Why do this now

Assessment is the broken slice of the learning and development value chain. Multiple choice became the default because it was the only format markable at scale, while everything richer still depends on scarce human markers who are slow, inconsistent and expensive. In South Africa the accredited assessor and moderator engine is visibly strained, and moderation backlogs stall accreditation outcomes for training providers.

AI marking of qualitative answers against a proper rubric, distributed as a simple link, does not exist as a product today. It will, and soon. This MVP proves it first, with real learners, for a fixed R150 000.

The rubric studio

CONSISTENT · DEFENSIBLE · YOURS

AI helps the assessment creator turn a question and a learning outcome into criteria, weightings and a model answer. The creator refines and approves; marking only ever runs against an approved, versioned rubric. That is what makes every mark consistent and explainable.

- Rubrics drafted in minutes, not meetings
- Every mark traceable to a rubric version

The marking engine

INSTANT · EXPLAINABLE · MEASURED

Typed answers are marked the moment they arrive: a score and written feedback per rubric criterion, with a consistency tolerance the platform must meet. Agreement is measured against a human-marked benchmark set of at least 50 answers, so trust is evidence, not assertion.

- Marking turnaround from weeks to seconds
- A published agreement rate to show sceptical buyers

The compliance unlock

CETA / QCTO ALIGNED · HUMAN IN THE LOOP

The frameworks do not prohibit AI-assisted marking; they require an accredited assessor and moderator overseeing outcomes. The MVP is built around exactly that: AI does the volume, low-confidence marks route to a moderation queue, and every override lands in an immutable audit trail.

- Providers retire the marking bottleneck
- An evidence pack an external verifier can read

Distribution without integration

ANY DEVICE · ANY LMS · NO LOGIN

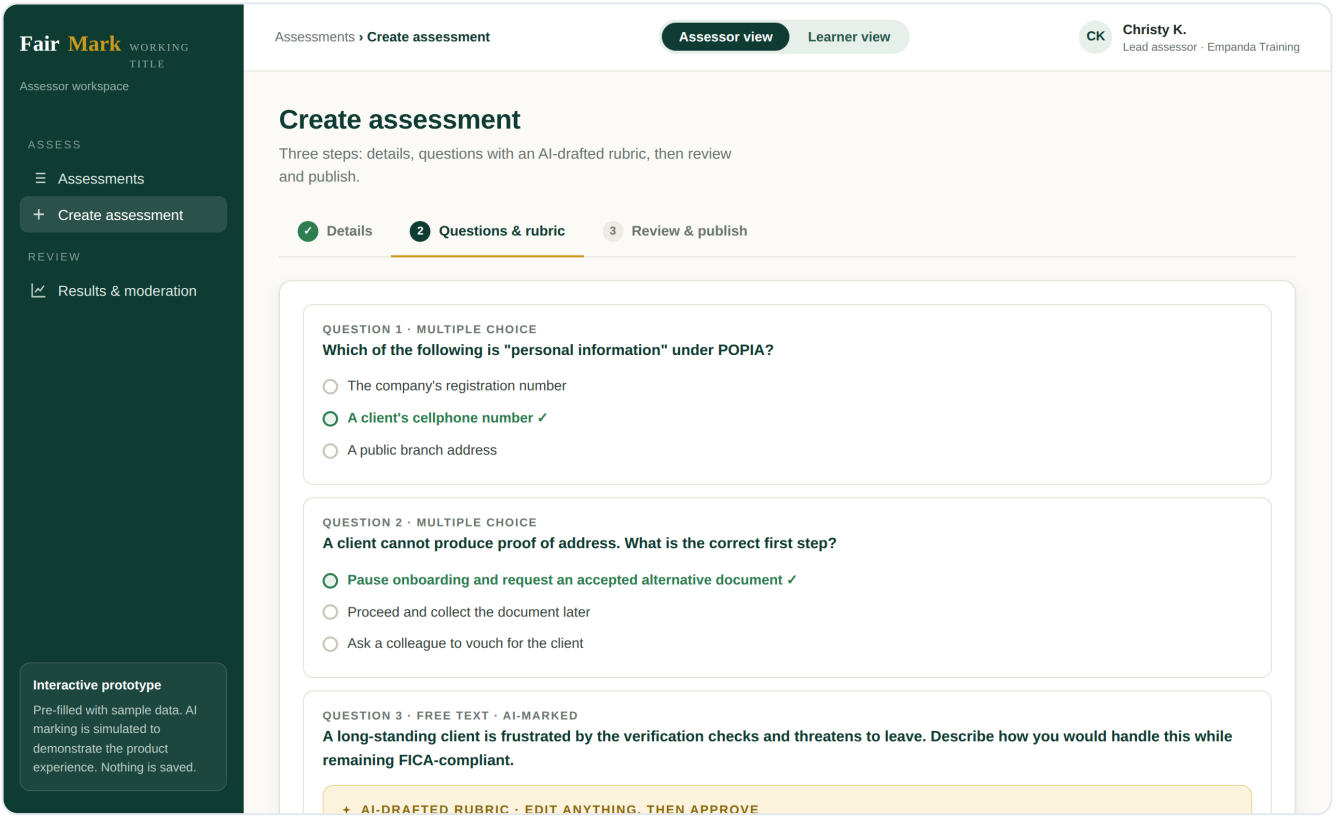
Each assessment is a tokenised link. It opens on a flagship laptop or a low-end Android phone from a WhatsApp message, embeds in an LMS page, and needs no account, app or integration project. Buyers can adopt it without asking IT for anything.

- Zero integration cost for the pilot client
- Learners complete assessments where they already are

02 What we will build

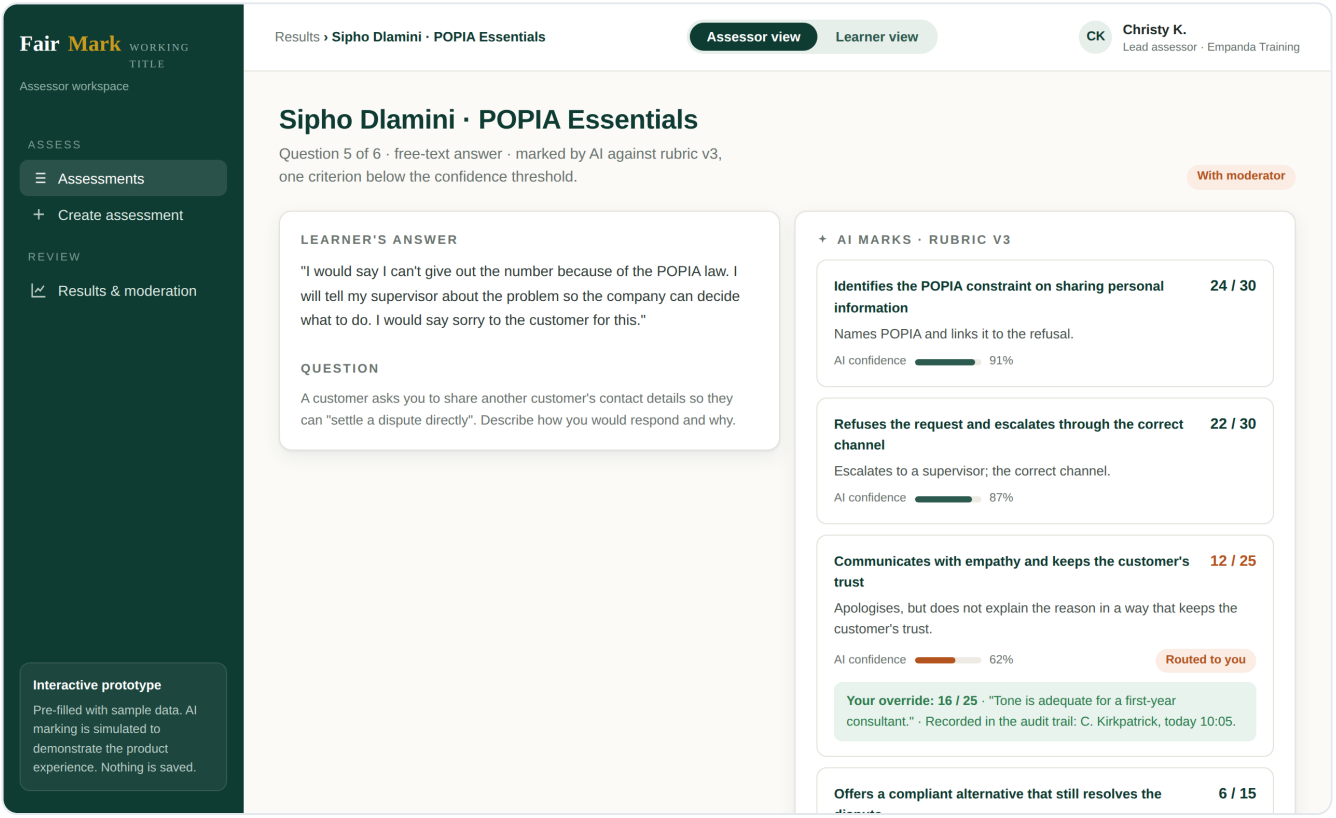
Nine backlog items deliver the full concept-proving loop for one pilot organisation: author, AI-draft the rubric, deliver by link, mark by AI, moderate by exception, release and export.

Rubric studio: AI drafts, the assessor approves



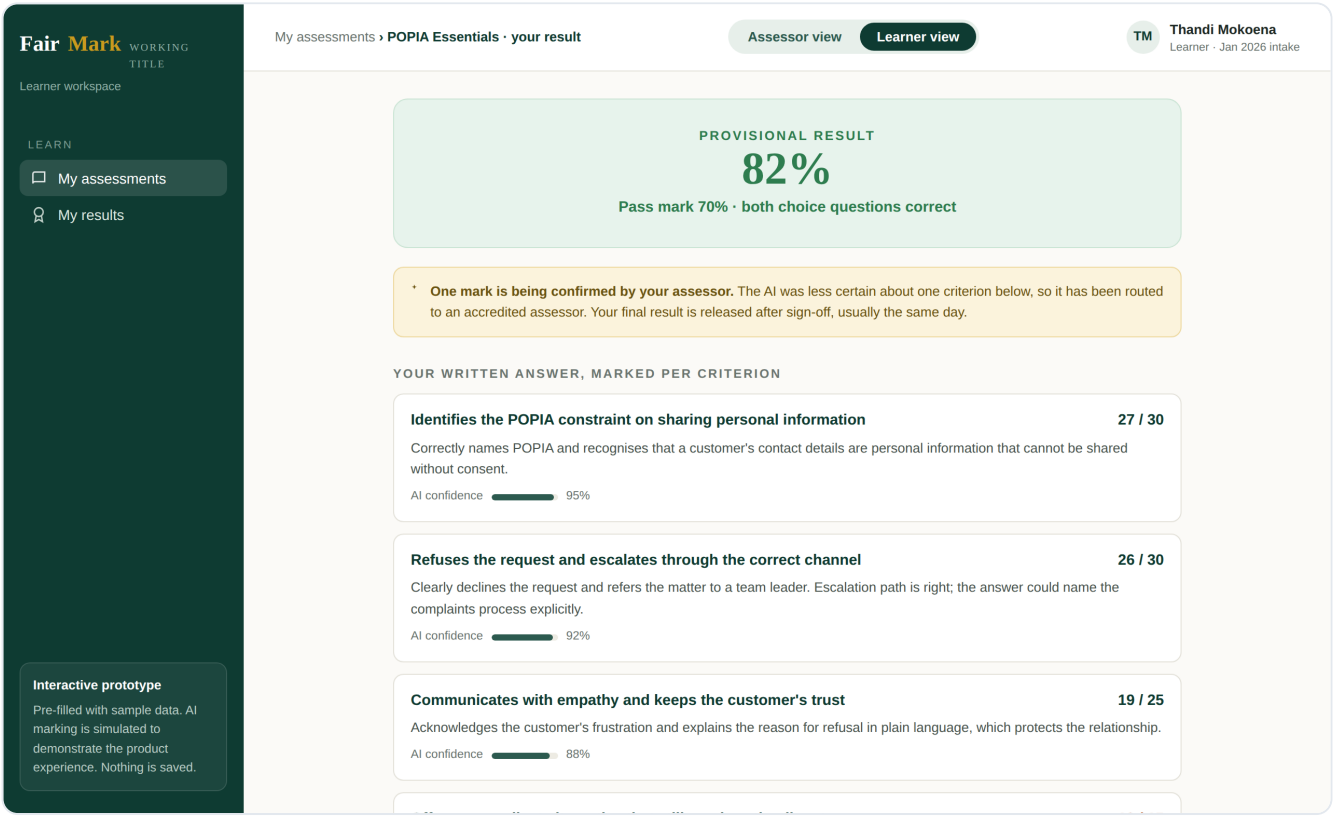
Prototype: the creator builds choice and free-text questions; AI drafts the rubric criteria, weights and model answer for the free-text question, and nothing marks until the assessor approves it.

Moderation: AI marks, an accredited human signs off



Prototype: a low-confidence criterion routed to the moderator, reviewed side by side with the learner's answer, overridden with a reason, and recorded in the audit trail before release.

The learner's result: instant, explained, trusted



Prototype: a provisional result seconds after submission, with a score and written feedback per rubric criterion and a clear note where the assessor is confirming a mark.

03 How it works

- **Author:** The creator builds an assessment under a topic, and AI drafts the marking rubric and model answer for every free-text question. The creator edits and approves; the version locks on publish.
- **Deliver:** Assigned learners receive a tokenised link that opens on any device with no login. Answers autosave, and a dropped connection resumes at the same question.
- **Mark:** AI marks each typed answer against the approved rubric: a score and short justification per criterion, checked for consistency and benchmarked against the human-marked answer set.
- **Moderate and release:** Low-confidence marks route to the accredited assessor, every override is recorded immutably, and results release only under the policy you set.

04 What's in and what's out

The MVP is deliberately smaller than the full platform vision: enough to prove trustworthy AI marking with a real cohort and support a raise, at a fixed price. Everything in the right column is planned, priced indicatively, and deferred on purpose.

In the MVP · R150 000 fixed

- One pilot organisation with creator, moderator and admin roles
- Assessment builder: multiple choice and free-text questions
- AI-drafted rubrics with model answers, versioned and assessor-approved
- Tokenised delivery link on any device, autosave and resume, basic iframe embed
- AI marking per rubric criterion with written feedback, benchmarked against your human-marked set of 50+ answers
- Moderation queue with override, sign-off and an immutable audit trail
- Results list with per-criterion detail and CSV export
- Pilot go-live with checklist, smoke test and rollback plan

Deliberately deferred to the scale roadmap · R325 000 indicative

- Multi-tenancy with per-organisation data isolation (the pilot data model is rebuilt for this)
- Voice answers, transcription and multilingual capture (MVP marks typed answers, verified for English)
- Proctoring: photo identity, timing, focus logging
- Tiered model routing, confidence auto-routing and cost telemetry
- Dashboards and standards-aligned accreditation reporting
- Results API, Marking-as-a-Service API, rubric API and QTI import
- Subscription billing with usage metering
- Documentation, formal UAT cycles and a hypercare window

These deferrals are what hold the MVP at R150 000: each one adds scale-readiness, not proof. The roadmap is re-baselined against pilot actuals before contracting.

05 Effort & sizing

Effort is expressed with our Small / Medium / Large sizing and a build-point scale, giving a single comparable measure of build size across the project.

Size	Effort	Build points
S Small	Half a day	1
M Medium	One day	2
L Large	Two days	3

Foundation and access

What it does	Size	Points
Production environment with CI/CD, secrets vault, daily backups and error alerting.	L	3
Creator, moderator and admin roles for the pilot organisation; learners join by tokenised link, no accounts.	M	2
Sub-total		**5**

Authoring and the rubric studio

What it does	Size	Points
Assessment and question authoring with locked publish versions; every attempt records the version it ran on.	L	3
Multiple choice with automatic scoring, single and multiple correct options.	S	1
AI drafts rubric criteria, weights, level descriptors and a model answer per free-text question.	L	3
Review, edit and approve flow with rubric versioning; unapproved rubrics cannot mark.	M	2
Sub-total		**9**

Learner delivery

What it does	Size	Points
Mobile-first assessment player opened from a tokenised link, verified on the agreed lean device set.	L	3
Autosave with resume after a dropped connection, plus a basic iframe embed.	S	1
Sub-total		**4**

AI marking and moderation

What it does	Size	Points
Per-criterion AI marking with written feedback and a consistency tolerance the platform must meet.	L	3
Marking queue with retries and full mark provenance: model, prompt version and rubric version.	M	2
Agreement measured against your human-marked benchmark set of at least 50 answers.	S	1
Moderation queue with side-by-side review and per-criterion override with a mandatory reason.	M	2
Immutable audit trail and a release policy gate on every result.	S	1
Sub-total		**9**

Results and go-live

What it does	Size	Points
Results list with per-criterion detail per learner.	S	1
CSV export matching on-screen filters.	S	1
Pilot go-live: checklist, smoke test and rollback plan.	S	1
Sub-total		**3**

Total build effort

Workstream	Sizes	Build points
Foundation and access	1 × L, 1 × M	**5**
Authoring and the rubric studio	2 × L, 1 × S, 1 × M	**9**
Learner delivery	1 × L, 1 × S	**4**
AI marking and moderation	1 × L, 2 × M, 2 × S	**9**
Results and go-live	3 × S	**3**
Total build points		**30**

Project activities

Alongside the build, delivery includes:

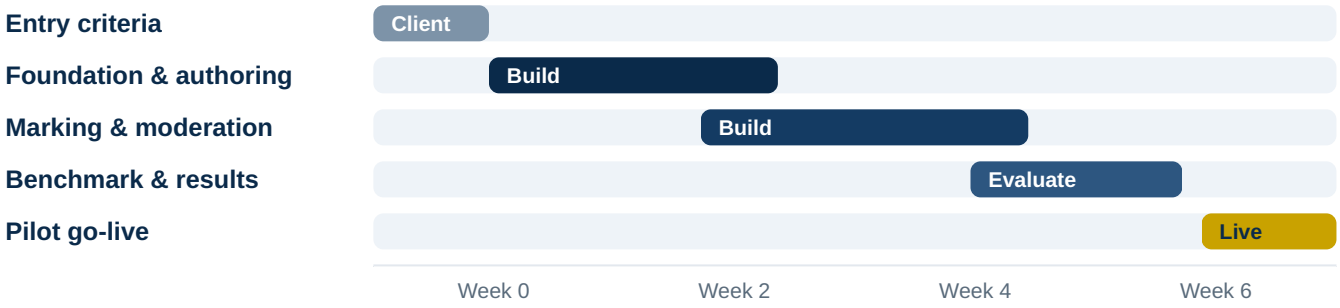
- **Benchmark evaluation** - the AI marking agreement rate is measured against your human-marked set before the pilot cohort sees a single mark, and reported to you.
- **Pilot support** - best-effort support through the pilot window; a formal hypercare window and UAT cycles sit on the scale roadmap.
- **Delivery management** - kept deliberately lean for a concept-proving build: weekly progress notes and a single point of contact.

06 Delivery plan

The MVP fits a focused six-week window from entry criteria to pilot go-live, followed by the pilot cohort running live. The scale roadmap is re-baselined against pilot actuals and typically funded from the raise the MVP enables.

Entry criteria	Build sprints	Pilot go-live	
BEFORE SPRINT 1	WEEKS 1 TO 5	WEEK 6	POST-RAISE
◆ Human-marked benchmark set of 50+ answers ◆ Data residency and POPIA hosting decision ◆ Product name and brand assets	◆ Foundation, authoring and rubric studio ◆ Delivery link, marking engine and moderation ◆ Benchmark evaluation and results	◆ Checklist, smoke test, rollback plan ◆ Pilot cohort live with real learners	◆ Multi-tenancy, dashboards, voice, proctoring ◆ Results and Marking-as-a-Service APIs, billing

Indicative timeline



07 Key assumptions

- The MVP is a concept-proving build, deliberately not scale-ready: one pilot organisation (single tenant), lean environments, single-model AI marking, typed answers, best-effort pilot support. The pilot data model is rebuilt for multi-tenancy in the scale phase.
- Custom engineering build (Next.js, Supabase, Azure), greenfield, with AI marking and rubric drafting on the Claude API. Marking quality is formally verified for English in the MVP.
- Voice answers, multilingual capture, proctoring, dashboards, the results and Marking-as-a-Service APIs and billing all sit on the scale roadmap (indicative R325 000, plus or minus 25 percent, re-baselined after the pilot).
- Christy's accredited assessor and moderator capability provides the compliance oversight; the platform provides the workflow, evidence and audit trail, not the accreditation itself.
- Pilot scale is up to roughly 100 concurrent learners on the agreed lean device set.
- Entry criteria are unbilled client preconditions and each must close before the relevant sprint starts; they protect the fixed price without padding it.

08 Investment

A single fixed price for the concept-proving MVP. The scale roadmap is indicative and contracted separately once pilot evidence is in.

AI Assessment and Marking Platform · concept-proving MVP

Fixed price for the complete MVP scope in this proposal · 30 build points

R150 000

EXCL. VAT

WHAT IS INCLUDED

- Authoring with the AI rubric studio
- AI marking benchmarked against your human-marked set
- Results, CSV export and pilot go-live
- Tokenised delivery link on any device
- Moderation queue with immutable audit trail

Billing follows riivo's effort-weighted milestones: 20 percent on story sign-off, 60 percent on development completion and 20 percent on client QC per sprint. The scale roadmap is indicative at R325 000 (65 build points, plus or minus 25 percent) and is re-baselined against pilot actuals.

09 Roles & approval

Role	Who
Driver	riivo delivery team
Approver	Christy Kirkpatrick, Empanda
Responsible	riivo build team; Christy's accredited assessor for moderation
Informed	Pilot organisation stakeholders

→ **Next step:** confirm the MVP scope and entry criteria, and riivo schedules Sprint 1. The pilot can be live with real learners inside six weeks of the benchmark set arriving.